Meta-analysis of computer-aided psychotherapy: Problems and Partial Solutions

Marks, I.M.; Cuijpers, P.; Cavanagh, K; van Straten, A.; Gega, L.; Andersson, G.

published in
Cognitive Behaviour Therapy
2009

DOI (link to publisher)
10.1080/16506070802675239

document version
Publisher's PDF, also known as Version of record

Cognitive Behaviour Therapy
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713926011

Meta-Analysis of Computer-Aided Psychotherapy: Problems and Partial Solutions

Isaac M. Marks[a]; Pim Cuijpers[b]; Kate Cavanagh[c]; Annemieke van Straten[b]; Lina Gega[d]; Gerhard Andersson[e]
[a] King's College London, Institute of Psychiatry, London, UK [b] Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands [c] School of Psychology, University of Newcastle, Newcastle, and Northumberland, Tyne and Wear NHS Trust, Newcastle upon Tyne, UK [d] School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK [e] Department of Behavioural Sciences and Learning, Swedish Institute for Disability Research, Linköping University, Linköping, and Department of Clinical Neuroscience, Psychiatry Section, Karolinska Institute, Stockholm, Sweden

To cite this Article Marks, Isaac M. , Cuijpers, Pim , Cavanagh, Kate , van Straten, Annemieke , Gega, Lina and Andersson, Gerhard(2009) 'Meta-Analysis of Computer-Aided Psychotherapy: Problems and Partial Solutions', Cognitive Behaviour Therapy, 38: 2, 83 — 90
To link to this Article: DOI: 10.1080/16506070802675239
URL: http://dx.doi.org/10.1080/16506070802675239

Cognitive Behaviour Therapy Vol 38, No 2, pp. 83–90, 2009

Meta-Analysis of Computer-Aided Psychotherapy: Problems and Partial Solutions

Isaac M. Marks[1], Pim Cuijpers[2], Kate Cavanagh[3], Annemieke van Straten[2], Lina Gega[4] and Gerhard Andersson[5]

[1]*King's College London, Institute of Psychiatry, London, UK;* [2]*Department of Clinical Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands;* [3]*School of Psychology, University of Newcastle, Newcastle, and Northumberland, Tyne and Wear NHS Trust, Newcastle upon Tyne, UK;* [4]*School of Medicine, Health Policy and Practice, University of East Anglia, Norwich, UK; and* [5]*Department of Behavioural Sciences and Learning, Swedish Institute for Disability Research, Linköping University, Linköping, and Department of Clinical Neuroscience, Psychiatry Section, Karolinska Institute, Stockholm, Sweden*

Abstract. Meta-analysis can be valuable if it heeds its originators' caution that intimate communing with the data is essential. A critique of the authors' own meta-analysis shows that the danger of overly broad conclusions could be reduced by attention to specificities and awareness of potentially hidden sources of variance. Conclusions from even good meta-analyses are best placed in perspective, along with naturalistic reviews, open studies, and even anecdotes to yield a fair picture of what computer-aided psychotherapy or any other treatment can achieve under varying conditions. The most realistic picture comes from zooming in and out and melding meta-analyses with further types of evidence. *Key words: meta-analysis critique; computer-aided psychotherapy; self-help.*

Received 21 November, 2008; Accepted 20 December, 2008

Correspondence address: Isaac Marks, PhD, MD, King's College London, Institute of Psychiatry, 43 Dulwich Common, London SE217EU, UK. Tel: +44 208 299 4130. E-mail: i.marks@iop.kcl.ac.uk

The plethora of randomised controlled trials (RCTs) in health care has stimulated researchers to devise ways of summarising outcomes across RCTs of particular treatments for particular problems by performing meta-analyses (MAs). MAs use quantitative methods to combine results from different studies to give a single estimate of effect and to help clinicians, researchers, and health care funders see the wood for the trees.

Methods to design, run, and analyse RCTs and perform MAs for outcome analysis across RCTs have improved iteratively since their start decades ago (Smith & Glass, 1977). MAs, a form of systematic review, have been of undoubted benefit in speeding the sifting of masses of health care evidence. MAs have been boosted by Cochrane reviews (Cochrane Collaboration, 2008), which claim to be reliable, independent, and comprehensive—the "gold standard" of evidence-based health care to influence decision-making by health carers, researchers, and policy-makers.

Busy people rely on the kitemarks of MAs and of Cochrane reviews to see trends at a glance. It is thus timely to audit such audits. How comprehensive and accurate are MAs? How carefully can their conclusions reflect the complexities of comparisons across RCTs, which may differ widely in the questions they ask and answers given based on diverse designs, measures and types of analyses, and how generalisable can the conclusions of MAs across such varied studies be to everyday practice in health care? Some pitfalls in MAs were noted at their start by Smith and Glass

 ISSN 1650-6073
DOI 10.1080/16506070802675239

(1977) and more recently by Rosenthal and DiMatteo (2001) and Sensky (2005).

Important differences can all too easily be obscured by computer-programmed calculations to several decimal points of standardized mean effect size (ES), random- and fixed-effects models, Q and I^2 statistics of heterogeneity, tests of publication bias, relative risks, odds ratios, confidence intervals, concealment of randomisation, dropouts at each stage, and testing of rater blindness. Such statistics are needed but might still ignore vital details hidden in a short phrase in a paper or unreported but emergent only on chatting with researchers, who may have regarded such issues as being too trivial to report, an opinion presumably shared with the referees and editors of the journals, even high-impact ones, who published the results. Such details can obscure problems in interpreting MA evaluations of health care. This comment is no criticism of authors, referees, or editors. Potentially important but unreported issues may emerge only with the wisdom of hindsight long after publication, a point applying as much to our own studies as to those of other researchers.

The present report critiques our own MA of computer-aided psychotherapy (CP) for anxiety disorders in this issue (Cuijpers et al., 2009) to illustrate problems that are common in MAs and need hard work to solve even partially. The devil is in the detail. We ask how MAs might be made more transparent regarding how they choose the RCTs they include, the subgroups they analyse, and how far their findings can be generalised when upscaling implementation in widely differing health care systems. We find that conclusions can only be very tentative from even the most transparent and sophisticated MAs that are possible. Combining the results of MA with insights from narrative reviews (e.g. Marks, Cavanagh, & Gega, 2007) may allow a more balanced appreciation of the role of CP in health care, although this will not completely solve the problems described here. Our MA asked what the effects are in (1) anxiety disorders of (2) CP compared with (3) a contrasting (control, inactive) non-CP condition. The question seems simple enough, but the answer is not as a result of diversities that can trap the unwary.

Diversity across anxiety disorders

The term *anxiety disorder* is a movable feast that is not always easy to pin down. The *International Classification of Diseases* (10th edition [*ICD-10*]) and *Diagnostic and Statistical Manual of Mental Disorders* (fourth edition; *DSM-IV*]) categories can be incongruent (e.g. obsessive–compulsive disorder [OCD] is not an anxiety disorder in *ICD-10* but is in *DSM-IV*, while agoraphobic avoidance is excluded from the panic disorder category more in *ICD-10* than in *DSM-IV*. Agoraphobic avoidance with or without panic leads to more disability than does panic *per se*, so *ICD-10* and *DSM-IV* panic disorder samples may differ in degree of disability, whereas *DSM-IV* pure panic disorder samples without agoraphobia are less disabled than *DSM-IV* patients who have both panic and agoraphobic avoidance. The most common anxiety problems in primary care are probably generalised anxiety disorder and mixed anxiety/depression and are loose-knit diagnoses. Different anxiety disorder samples vary in, among others, chronicity, degree of disability, and course and in response to treatment, which itself may partly depend on the foregoing. MAs can try to handle this issue by including only RCTs of CP that concern just one *ICD* or *DSM* category, but at present this cuts the number of RCTs qualifying for inclusion drastically, thereby limiting conclusions.

Even if the MA concerns RCTs of patients with a clearly defined anxiety disorder, another source of variance is how the patients were sought and screened and what proportions were self- or general practitioner (GP) referrals or already on a wait list for face-to-face psychotherapy from mental health care professionals. The more filters they pass through to get into an RCT, the more likely they are to be more chronic and severe and perhaps more difficult to help. They might also be less motivated if they have been discouraged from doing CP by a referring mental health professional who worries about losing customers. Compared with referrals from mental health professionals, GP and self-referrals complied more with CP instructions and improved more (Mataix-Cols, Cameron, Gega, Kenwright, & Marks, 2006). As with single

diagnoses, MAs will currently find too few RCTs of CP with enough data about sources of referral to allow reliable subgroup analysis of this issue.

Diversity across CP

Different modes of delivering CP

Working out a mean ES of CP is about as useful as working out a mean ES of pharmacotherapy, covering a huge range of compounds and ways of getting them onto or into the body. Just as a vast number of medications might be delivered as an ointment, powder, or drops on the skin, eyes, or ears, inhaled as a spray, put under the tongue, swallowed as a tablet, capsule, powder, or wafer, injected subcutaneously, intramuscularly, or intravenously, or inserted as an anal or vaginal pessary, so CP systems can be delivered on a range of devices such as standalone or Internet-linked computers, PCs, palmtops, phone-interactive voice response, CD-ROMS, DVDs, cell phones, and virtual reality devices (Marks et al., 2007). Each of these modes of delivery has its own pros and cons. Lumping all RCTs into a single MA risks missing potentially crucial differences in effectiveness, with too few RCTs in given subgroups for meaningful analyses.

Differing contents and styles of CP systems

Mode of delivery aside, CP systems for anxiety disorders vary widely in their content and style. They may facilitate homework between face-to-face CBT sessions, give respiratory feedback, or aid vicarious in-session exposure in a clinic or direct self-exposure outside the clinic, writing exposure, or cognitive restructuring. In style, they may be page-turners (bibliotherapy on a screen), with tests like a college course suitable for educated users, may just show pictures on a screen, or may be more interactive simulations of a series of clinical sessions. RCTs rarely report duration and pacing of system use. These variables are hard to control, especially with CP used on the Internet at home or elsewhere outside a clinic. Users of given CP systems differ in the speed with which they can digest and implement CP guidance. Some may use a system for long sessions but infrequently over a long period, whereas others may use it often in brief sessions that can add up to a shorter total time and period of use. Such variables may affect the support that users get from a human if that is given regularly (e.g. once a week over the total period of system use) rather than when the user gets to a particular point in the system.

Differing durations and types of human contact during CP

CP, unlike the withdrawal of cash from an automatic-teller machine, is rarely completed with no human contact at all by e-mail, phone, or face-to-face at any stage from first screen to end of follow-up. CP is thus generally "CP plus some human contact, which varies vastly in duration and type" across RCTs. To be meaningful, an MA has to dissect these carefully, but this is easier said than done. Let us take examples from Marks et al. (2007). An RCT report of *Interapy* (Lange et al., 2003) did not mention therapist time, but, in fact, *Interapy* requires a mean of 14 hr of therapist support by e-mail (Ruwaard and Lange, personal communication, June 30, 2008 and July 9, 2008), whereas in the RCTs of Hassan (1992) and Bornas, Tortella-Feliu, Llabres, and Fullana (2001), a therapist sat with the patient throughout CP sessions. In these RCTs, a therapist spent as much time with CP patients as is usual during face-to-face exposure therapy in the United Kingdom, so by that the UK criterion the CP saved no therapist time or cost, a vital issue for cost-effectiveness. Palmtop and virtual reality systems had much therapist contact. Our MA of CP (Cuijpers et al., 2009, this issue) found the highest ESs of all in the Hassan (1992) and the Bornas et al. (2001) RCTs and a high ES in the Lange RCT; in those studies, however, CP was an add-on to a usual duration of therapist contact in the United Kingdom, even if that contact was just either sitting passively with the patient during CP or reading patients' e-mails and replying by e-mail.

At the other extreme of duration is zero therapist time with CP users, as in a multicentre RCT of patients using *BTSteps* at home (Greist et al., 2002). However, in that RCT, the multicentre ethical committees insisted on patients in each arm attending a clinic four times in the course of the RCT to see a clinician for brief safety checks, not therapy. Such clinic contact did not improve

non-CP patients, but perhaps it enhanced compliance with CP beyond that seen with CP under everyday conditions. MAs of RCTs asking patients to attend a clinic for whatever reason or to make repeated outcome ratings are unlikely to yield reliable conclusions about the outcome of CP used at home or in a library under field conditions. Casual unmonitored registrants at CP websites have far higher dropout rates than participants in the demanding conditions of RCTs of CP on the Internet (Eysenbach, 2005), one example being community users of the *MoodGym* program (Christensen, Griffiths, Groves, & Korten, 2006), with huge attrition rates.

In a meta-regression analysis, our MA found that ES rose with longer therapist contact. It could not determine whether this rise was directly proportional to longer duration of CP use, nor could it test dose–response issues adequately (e.g. whether just 5 min of instructive, encouraging contact once only or a few times might yield a higher ES than much longer contact) or whether effects differed according to content of the contact or level of expertise of the contact provider (i.e. a highly-trained professional or a very briefly-trained lay person). Nor could our MA test the interaction of such variables with different CP systems and different anxiety disorders. Our MA revealed an RCT that found better outcome with CP if users had scheduled rather than unscheduled phone support by a professional (Kenwright, Marks, Graham, Franses, & Mataix-Cols, 2005), but excluded it from analysis because both its arms involved CP (see below). Nor could our MA find enough RCTs for subgroups to judge whether contact was better face-to-face or by phone, e-mail, SMS text message, or a combination of these. This means an MA definition of heterogeneity in terms of ES alone is too simple, because studies yielding similar ESs were nevertheless heterogeneous in important other respects (e.g. reduction in therapist time varying from 0% to 100% or varying educational level of participants across studies). Such heterogeneities limit the comparability of, and conclusions that can safely be drawn from, MAs of RCTs.

RCTs are said to be a gold standard of evidence. Appropriately designed RCTs can indeed give good answers to carefully formulated questions under tightly regulated conditions. However, RCTs have limited ability to judge performance under less tightly-regulated conditions in routine care. This problem is inevitable with MAs, too, because they only analyse RCTs, not field studies of dissemination. Some dissemination issues might be tested by doing cluster RCTs, in which the units of randomisation to CP versus non-CP are separate geographical areas/institutions/other entities rather than individual patients. Unfortunately, such RCTs are very rare and need even more work and funding than RCTs that randomise individuals, so MAs cannot find enough to include. Moreover, in a cluster RCT, the non-CP arm would require as much attention-placebo and regulation as the CP arm, so the non-CP arm would no longer represent routine conditions for dissemination, even if patients are recruited in routine care and the therapists work in routine care. This uncertainty principle applying to RCTs applies to MAs of RCTs as well.

Diversity across control non-CP contrast conditions

RCTs are the raw material for MAs. A common belief is that RCTs can have an ideal control (contrast) group. In fact, any control group can only advance knowledge to the extent that it can answer the particular questions posed by the RCT. It is vital to scrutinise the potentially therapeutic ingredients in each group compared. For example, many RCTs compared CP plus human support with a wait-list (delayed-treatment) or treatment-as-usual (TAU) condition and judged CP to be superior. That superiority might reflect the greater activity and attention accompanying CP rather than the CP *per se*, in which case the CP would be redundant, like any other treatment that was no better than a comparable attention-placebo condition.

Such attention-placebo confounds cannot be compensated by the large sample sizes, masked randomisation, blind ratings, low dropout rates, and rigorous statistics that get good grades in systematic reviews. Different expectancies across groups could be reduced by wait-list or TAU patients being told that they may improve just with the passage of time, although they are rarely told this.

Wait-list and TAU contrast groups try to answer the question, "Does introducing CP improve the existing situation?" but cannot distinguish CP's effect from its accompanying attention-placebo effect unless the wait-list and TAU groups, too, have an added attention-placebo component; but then the RCT is no longer testing an ordinary wait-list or TAU condition. A regulatory body would not approve any drug or psychological treatment from results based on an RCT comparing it with a wait-list or TAU condition without an attention-placebo giving a comparable expectation of improvement. The same logic applies to CP trials. This questions our MA's inclusion of ESs from RCTs using wait-list or TAU contrast conditions. An active attention-placebo for phobias and obsessive–compulsive disorder (OCD) could, for example, be relaxation sessions without exposure + relaxation homework without exposure + relaxation diaries without exposure. Patients find such relaxation acceptable and tend to drop out less from it than from CP proper, even though relaxation without exposure is known to be ineffective for phobias and OCD, unlike applied relaxation, which does contain exposure (Öst, 1987). A second acceptable attention-placebo for anxiety disorders could be free-association sessions with matching homework and diaries excluding exposure. A third is antiexposure (Marks, 1987). Our MA's subgroup analysis found no effect on ES of type of control (wait-list vs. relaxation), but there were only two studies of relaxation without exposure, and "other" may be too broad a category to be meaningful, so this question remains unresolved.

Another non-CP issue concerns the definition of "wait list." Our MA excluded the study by Klein, Richards, and Austin (2006), because its wait-list group not only had access to an information website but also weekly contact in which they were asked whether they did their registration exercises. This further shows the difficulties in deciding which studies to include. "Wait list" can describe somewhat differing conditions across different authors, and it is not known how these may affect outcome.

MAs can be bedeviled by surprising results because of current lack of knowledge about which nonexposure psychotherapy procedures are good attention-placebo controls (contrasts) that reliably do not improve anxiety disorders. A case in point is our MA's legitimate exclusion of an RCT of CP guiding self-exposure for phobia/panic (Schneider, Mataix-Cols, Marks, & Bachofen, 2005) because its description of the control condition appeared, in fact, to contain active elements. However, its intended control attention-placebo condition, "managing anxiety," included relaxation, problem-solving, physical, and breathing exercises and distraction but excluded both exposure and systematic cognitive restructuring. Some serious cognitive behaviour therapists would not call this contrast condition CBT, yet it turned out to work unexpectedly well by posttreatment, although less so at follow-up. Our MA thus includes an RCT (Marks, Kenwright, McDonough, Whittaker, & Mataix-Cols, 2004) comparing CP with a computer-aided relaxation-placebo, which was ineffective at both posttreatment and follow-up, but excludes another RCT (Schneider et al., 2005) comparing similar CP with a different computer-aided condition that was intended to be an inactive attention-placebo throughout but turned out to be less effective than CP only at follow-up. Was the Schneider et al. (2005) RCT controlled for non-CP based on results at follow-up but not for non-CP based on results at posttreatment? A treatment arm intended to be an inactive attention placebo turned out to be surprisingly active in the short term. Related uncertainties about exclusion criteria arise with other studies (Fraser, Kirkby, Daniels, Gilroy, & Montgomery, 2001; Newman, Consoli, & Taylor, 1997). There is no clear answer. MAs, including our own, are inevitably arbitrary in some ways even when the meta-analyst tries to specify clear-cut criteria for including and excluding studies.

Clearly, complex conundrums confounded our choice of non-CP studies for inclusion in MAs. The choice can be moot however hard we strive for objectivity. The criterion for inclusion of CP versus a non-CP control can be too fuzzy to decide. Any study that contrasts an experimental treatment with a contrasting condition is, in a sense, controlled. For example, the Schneider et al. (2005) RCT controlled for presence of active exposure and cognitive restructuring instructions during computer guidance and so had a non-CP control. The Kenwright et al. (2005) RCT controlled for type of human contact during CP but had no non-CP control.

Diversity of outcome across measures

ESs can vary considerably across different measures within a given RCT, and simply pooling them is not always a satisfactory solution. For example, in a non-CP RCT of panic/agoraphobia (Marks et al., 1993), panic improved hugely with placebo but phobic avoidance and work/social adjustment did not. Target-phobia measures improve more with CBT, and earlier, than global-phobia measures, which, in turn, improve earlier than work/social adjustment measures. In such cases, one has to qualify outcomes in MAs regarding which measures improved and when.

Limitations of MAs

MAs lose much by excluding non-RCT evidence, although meta-analyses of open pre–post studies are possible but are not usually held in high regard. No kind of evaluation is ideal for all purposes. Sticking solely to a supposed gold standard of RCTs may neglect potentially crucial issues raised by anecdotes, case studies, and open trials. A few anecdotes make the point. First, in an early three-arm RCT of CP (Ghosh, Marks, & Carr, 1988) that compared self-exposure guidance given by CP, by a self-help book, or by face-to-face therapy, two patients who had been randomised to the book arm turned out to have bought that book in a shop but never opened it until they entered the trial and were asked by a psychiatrist to read the book and follow its guidance, which they did and then improved. The RCT's self-help book arm turned out to be testing not the value *per se* of a self-help book bought casually in a shop, but of that book recommended strongly by a health caregiver in a trial where the patients were asked to return to a clinic at intervals for assessment. How does one do an RCT to test the value of casually buying the book either with or without a professional recommendation? No RCTs means no MAs. Guidance on this kind of issue may have to rest on anecdotal evidence or perhaps epidemiological studies on the uptake and outcome of self-help materials in the general public.

Two more stories concern CP's confidentiality. During recruitment for CP trials, many patients said they valued CP's confidentiality and would not enter the trial if their GP had to be informed of the fact, and one said he would not dare enter the door of a CP clinic lest others outside saw him going there. Second, when a nurse asked if he could watch a patient doing CP at a computer screen in a clinic, the patient said, "Sure, just stand behind me." As certain information unfolded on the screen, the nurse exclaimed, "You didn't tell us that!," whereupon the patient turned and said, "Of course not; it's confidential." "Confidential" here implied that the patient would not reveal that information face-to-face but did not mind staff learning about it less directly, a subtlety rarely recognised in discussions of "confidential" although implied when patients in psychoanalytic sessions lie on a couch facing away from the analyst or when Catholics are screened from their priest during confession. Again, RCTs, and hence MAs, to tease out such issues are hard to devise.

RCTs and hence MAs are also hard put to address vital aspects of health care apart from efficacy (e.g. patient choice and preference). It is difficult enough for professionals to examine the value and convenience of different types of CP and their delivery. The snags are greater still for the average person on the street beset even more by many marketing variables, including the reservations of professionals who might mistakenly fear being supplanted by computers and of funders of health care. When RCT evidence convinced the UK National Institute for Health and Clinical Excellence (2006) to recommend that every primary care trust in England and Wales should use one CP system to manage depression and another to manage phobia/panic, many barriers to optimum implementation were encountered, including professionals' resistance, funding pressures, organisational rigidities, and contending self-help approaches.

There is no free lunch in health care. MAs can be a major step forward if they heed the original caution of Smith and Glass (1977): "Intimate communing with the data … is an essential … requirement" (p. 756). More intimate communing with the data allows a better grasp of the problems of MAs but cannot eliminate them. Awareness of the limitations of blunt MAs plus sharper attention to hidden sources of variance could allow more realistically qualified conclusions to be drawn.

Questions asked by an MA become more useful when they move from the form "Is CP effective for anxiety disorders?" to: "Compared with contrast-treatment condition A, how much did CP system B improve sufferers from anxiety-disorder C referred in manner D and screened by method E who did CP with F amount of human support delivered in mode G by supporters who were trained by method H?" Even these specificities will not necessarily tell us how much improvement will follow from disseminating the use of system B into hugely varying field conditions over a whole country. The same is true for RCTs (and thus MAs) of any treatment, be it a CP system, a drug, acupuncture, mindfulness meditation, or whatever. "Does that treatment work?" becomes a more meaningful question when refined into: "Does treatment V improve clinical problem W if applied by method X for Y months with surveillance every Z months?" Such nuancing can counter a cynic's quip that "meta-analysis is to analysis what metaphysics is to physics." Unfortunately, such nuancing sharply cuts the numbers of RCTs available for MA study of each issue.

Conclusions

Search for a holy grail to evaluate therapy will be in vain. Different strategies to review therapies resemble ways of viewing the world using a microscope or Google Earth. Zooming in reveals detail but not the big picture, which may be crucial. Zooming out shows the overall shape of things but not potentially vital details. Reviews are most informative when they acknowledge their distance of view and flexibly zoom in and out, combining meta-analytic, narrative, and other approaches. Together, these give policymakers a more balanced guide to the best evidence available and enlighten researchers and practitioners closer to the ground on the subtleties of study design and outcomes that can move knowledge forward in different ways. Each strategy has its own strengths and weaknesses.

We all want a bottom line, a simple answer to "Does the treatment work?" plus a one-page executive summary supporting the answer. Because of uncertainties inherent in MAs, however, a bottom line from even the best MAs is not reliable unless placed in context. Meta-analyses yield important evidence to put in perspective along with naturalistic reviews, open studies, and even anecdotes to yield a realistic picture of what CP and any other treatment can achieve.

Acknowledgments

Isaac Marks shares intellectual property rights in the *BTSteps*, *FearFighter*, and *Cope* computer-aided psychotherapy systems. Kate Cavanagh is an occasional consultant to CCBT Ltd., which markets *FearFighter*, *BTSTeps*, and *Cope*, and to Ultrasis plc, which markets *Beating the Blues*.

References

Bornas, X., Tortella-Feliu, M., Llabres, J., & Fullana, M. A. (2001). Computer-assisted exposure treatment for flight phobia: Controlled study. *Psychotherapy Research*, *11*, 259–273.

Christensen, H., Griffiths, K., Groves, C., & Korten, A. (2006). Free range users and one hit wonders: Community users of an Internet-based cognitive behaviour therapy program. *Australian and New Zealand Journal of Psychiatry*, *40*, 59–62.

Cochrane Collaboration. (2008). *An introduction to Cochrane reviews and The Cochrane Library*. Retrieved September 25, 2008, from http://www.cochrane.org/reviews/clibintro.htm.

Cuijpers, P., Marks, I. M., Cavanagh, K., van Straten, A., Gega, L., & Andersson, G. (2009). Computer-aided psychotherapy for anxiety disorders: A meta-analytic review. *Cognitive Behaviour Therapy*, *38* (2), 66–82.

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research*, *7*(1), e11.

Fraser, J., Kirkby, K. C., Daniels, B. A., Gilroy, L. G., & Montgomery, I. M. (2001). Three versus six sessions of computer-aided vicarious exposure treatment for spider phobia. *Behaviour Change*, *18*, 213–224.

Ghosh, A., Marks, I. M., & Carr, A. (1988). Therapist contact and outcome of self-exposure treatment for phobias. *British Journal of Psychiatry*, *152*, 234–238.

Greist, J. H., Marks, I. M., Baer, L., Kobak, K. A., Wenzel, K. W., & Hirsch, M. J., et al. (2002). Behaviour therapy for obsessive compulsive disorder guided by a computer or by a clinician compared with relaxation as a control. *Journal of Clinical Psychiatry*, *63*, 138–145.

Hassan, A. A. M. (1992). *Comparison of computer-based symbolic modelling and conventional methods in treatment of spider phobia*. Unpublished doctoral dissertation, University of Leeds.

Kenwright, M., Marks, I. M., Graham, C., Franses, A., & Mataix-Cols, D. (2005). Brief scheduled phone support from a clinician to enhance computer-aided self-help for OCD: RCT. *Journal of Clinical Psychology*, *61*, 1499–1508.

Klein, B., Richards, J. C., & Austin, D. (2006). Efficacy of Internet therapy for panic disorder. *Journal of Behavior Therapy and Experimental Psychiatry*, *37*, 213–238.

Lange, A., Rietdijk, D., Hudcovicova, M., van den Ven, J.-P., Schrieken, B., & Emmelkamp, P. M. G. (2003). Interapy: RCT of the standardized treatment of posttraumatic stress through the Internet. *Journal of Consulting and Clinical Psychology*, *71*, 901–909.

Marks, I. M. (1987). *Fears, phobias and rituals: Panic, anxiety, and their disorders*. New York: Oxford University Press.

Marks, I. M., Cavanagh, K., & Gega, L. (2007). *Hands-on help: Computer-aided psychotherapy* (Maudsley Monograph No. 49). Hove, UK: Psychology Press.

Marks, I. M., Kenwright, M., McDonough, M., Whittaker, M., & Mataix-Cols, D. (2004). Saving clinicians' time by delegating routine aspects of therapy to a computer: A randomised controlled trial in phobia/panic disorder. *Psychological Medicine*, *34*, 9–18.

Marks, I. M., Swinson, R. P., Başoğlu, M., Kuch, K., Noshirvani, H., & O'Sullivan, G., et al. (1993). Alprazolam and exposure alone and combined in panic disorder with agoraphobia: A controlled study in London and Toronto. *British Journal of Psychiatry*, *162*, 776–787.

Mataix-Cols, D., Cameron, R., Gega, L., Kenwright, M., & Marks, I. M. (2006). Effect of referral source on outcome with CBT self-help. *Comprehensive Psychiatry*, *47*, 241–245.

National Institute for Health and Clinical Excellence. (2006). *Computerised cognitive behavioural therapy (CCBT) for the treatment of depression and anxiety: Technology Appraisal No. 97*. London: Author.

Newman, M. G., Consoli, A., & Taylor, C. B. (1997). Computers in assessment and CBT of clinical disorders: Anxiety as a case in point. *Behavior Therapy*, *28*, 211–235.

Öst, L.-G. (1987). Applied relaxation: Description of a coping technique and review of controlled studies. *Behaviour Research and Therapy*, *25*, 379–409.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*, 59–82.

Schneider, A. J., Mataix-Cols, D., Marks, I. M., & Bachofen, M. (2005). RCT in phobia/panic disorder of two forms of net-guided self-help, each with brief live phone support. *Psychotherapy and Psychosomatics*, *74*, 154–164.

Sensky, T. (2005). The effectiveness of cognitive therapy for schizophrenia: What can we learn from the meta-analyses? *Psychotherapy and Psychosomatics*, *74*, 131–135.

Smith, M., & Glass, G. (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, *32*, 752–760.